

Modelling of Baker's Yeast Production

Havva Boyacıoğlu^{1*}, Suna Ertunç¹, Hale Hapoğlu¹

¹Department of Chemical Engineering, Faculty of Engineering, Ankara University, Ankara, Turkey

Received: 04 May 2016 - Accepted: 08 August 2016

Abstract: In the present work, parametric models for the control of bioreactor temperature have been applied. Various order discrete time model parameters were evaluated theoretically and experimentally. Two types of input signals were used as external force to determine Auto Regressive Moving Average with Exogenous (ARMAX) model parameters with Recursive Least Square (RLS) parameter estimation algorithm. The third order ARMAX model is utilized, and compared with the second order one. Ternary and square disturbances are given to the cooling water flow rate which can be chosen as manipulating variable in closed loop cases. System response is monitored continuously and the model parameters are calculated. The models with experimentally identified parameters are compared with ones that their parameters are identified theoretically.

Keywords: Baker's Yeast, System Identification, *Saccharomyces cerevisiae* as a second metabolite source

1. Introduction

As *S. cerevisiae* investigations continue in research and development of food and drug need improvement, as well as biotechnological and genetic purposes. The literature on microbial genomic sequences highlighted the necessity for the production of desired metabolites. The important compounds *S. cerevisiae* have importance commercially as one of the most studied model organisms with large scale bioreactor operation. The control performance can be enable a considerable increase in the industrial application of Baker's Yeast. Studies have described *S. cerevisiae* production that there is great emphasize on pharmaceutical usage, the investigations related with the production is then encouraged to explore the data in order to obtain a better understanding of it [1, 2].

Baker's yeast production can be achieved by means of batch or fed-batch operation under aerobic conditions. Time varying behavior of bioprocesses exhibits complex, nonlinear behavior of which the modelling is very difficult [1].

Examination of the internal structure of bio-systems illustrates a complete case with unknown factors and unmeasurable variables [3]. According to requirements, order and accuracy definitions may be changed for bioprocess model applications. The modelling with parameters identification is one of the effective procedures to define the systems. Therefore two main part of it include choosing the model structures and evaluation of any certain model parameters. An acceptable approximation of the system should be achieved by utilizing the

*Corresponding Author E-mail: havvaboyacioglu@gmail.com

best estimates of model degree and all the unknown variables of operation. Considerable computation time must be also spent for parameter estimation procedure [4].

For process control applications, proposed system models were usually written in discrete-time domain. The input and output variables were sampled to identify process parameters, although the process is implemented in continuous-time nature, and represented by differential equations. Svoronos et al., (1981) reported a bilinear model based upon minimum-variance self-tuning rule. Their model effectiveness was examined by using the simulations. Hapoglu et al. (2001) utilized a Controlled Auto Regressive Integrated Moving Average (CARIMA) model and its parameters were identified with Bierman computation procedure in which data obtained by enforcing the system with a pseudo random binary sequence (PRBS). Akay et al. (2003) investigated parametric and non-parametric models which include the relationship between dissolved oxygen concentration and air flow rate in *S. cerevisiae* production medium. These models theoretical and experimental identification were realized. During process control application, the manipulated variable was inlet air flow rate. Controlled Auto Regressive Moving Average (CARMA) model parameters were evaluated by disturbing the system with various types of input signals and using Recursive Least Square (RLS) parameter estimation algorithm. The aeration step response was obtained to maintain the dynamic matrix and the non-parametric model [7].

The actual part of the system can be modelled with approximated structure and estimated parameters of it. During the control, the closed loop performance of the system highly depends on the real process and its model mismatch. Many models relate input-output in a non-linear and linear fashion, such as, Hammerstein type (Zhu and Seborg, 1994), Wiener type (Norquay et al., 1996), Non-linear state predictor (NSP) form (Wang et al., 2004), subspace-based linear type (Sotomayor et al., 2003), linear multivariable discrete-time model [12]. Akay et al. (2011) designed a batch bioprocess parametric model and optimized the model order with experimental transient behavior analysis. Experimental DO control of the medium was realized and compared with theoretical results [1].

In the present investigation, the models of Baker's yeast production in a batch process are obtained experimentally and theoretically. Various ARMAX type models are utilized and tested.

1.1. Discrete-time System Models

An ARMAX model in discrete-time domain for the single input single output system representation is utilized as following:

$$A(z^{-1})y(t) = z^{-k}B(z^{-1})u(t) + C(z^{-1})e(t) \quad (1)$$

Where k is system velocity lag, y(t) is the response of the system, u(t) represents the most effective input variable, A, B and C are polynomials.

$$A(z^{-1}) = 1 + a_1z^{-1} + a_2z^{-2} + \dots + a_{na}z^{-na} \quad (2)$$

$$B(z^{-1}) = b_0 + b_1z^{-1} + b_2z^{-2} + \dots + b_{nb}z^{-nb} \quad (3)$$

$$C(z^{-1}) = 1 + c_1z^{-1} + c_2z^{-2} + \dots + c_{nc}z^{-nc} \quad (4)$$

By using matrix notation, the system model can be rearranged as below:

$$y(t) = X^T(t)\theta(t) + \varepsilon(t) \quad (5)$$

Where X and θ vectors consist of the data and parameters.

$$X^T(t) = [y(t-1), \dots, y(t-na), u(t-1), \dots, u(t-nb-1), 1, e(t-1), \dots, e(t-nc)] \quad (6)$$

$$\theta(t) = [a_1, a_2, \dots, a_{na}, b_0, b_1, \dots, b_{nb}, c_1, \dots, c_{nc}] \quad (7)$$

1.2. Recursive Least Square Identification

The predicted and measured system responses differences are evaluated and sum of the squares of these is minimized as parameter identification criteria. This system identification technique (Soderström and Stoica, 1998) describe error prediction (see Eq.8) by means of model in Eq.5.

$$\varepsilon(t+1) = y(t+1) - X^T(t+1)\theta(t) \quad (8)$$

Recursive evaluation of Eq. 9 and Eq. 10 are realized until the stopping criterion is satisfied.

$$P(t+1) = P(t) - \frac{P(t)X(t+1)X^T(t+1)P(t)}{1 + X^T(t+1)P(t)X(t+1)} \quad (9)$$

$$\theta(t+1) = \theta(t) + P(t+1)X(t+1)\varepsilon(t+1) \quad (10)$$

In the cases studied, the model testing was achieved by using the integral of absolute error (IAE) and the integral square of error (ISE) criteria (see Eq. 11 and Eq. 12) and parameter estimation error norm (PEEN) criteria which is given in Eq. 13.

$$IAE = \sum_{t=1}^N |\varepsilon(t)| \quad (11)$$

$$ISE = \sum_{t=1}^N |\varepsilon(t)|^2 \quad (12)$$

$$PEEN = \frac{\sum_{i=1}^n |y(t) - X^T(t)\theta(t-1)|}{\sum_{i=1}^n |y(t)|} * 100 \quad (13)$$

2. Materials and Methods

The Northern Regional Research Centre, ARS Culture Collection (Peoria, IL, USA) microorganism NRRL-Y-567 was used. The growth medium and scaling details are same as the previous works [1, 7].

A 2 L jacketed bioreactor was used and equipped with the following: an oxygen sensor, a pH sensor, a thermocouple, a 4 bladed turbine type impeller, an immersed heater, air supplier, a pump for cooling water, a circulator, a V/I converter, a rotameter, an I/P transducer, a microbiological filter and an on-line computer control system. Experiments were carried out at the optimum growth conditions of temperature 32°C and pH 5. Agitation rate was maintained at 600 rpm and cooling water was passed through the jacket at 21°C. Foam

formation was prevented by adding antifoam bioreactor at certain time intervals and microorganism growth was observed by using UV spectrophotometer analyses. An on-line computer with an I/O module receives DO, pH and temperature signals at every sampling instants. The on-line computer with VISIDAQ data acquisition and programming package in which cooling water flow rate was adjusted continuously in the jacket was used. The model identification was achieved in MATLAB.

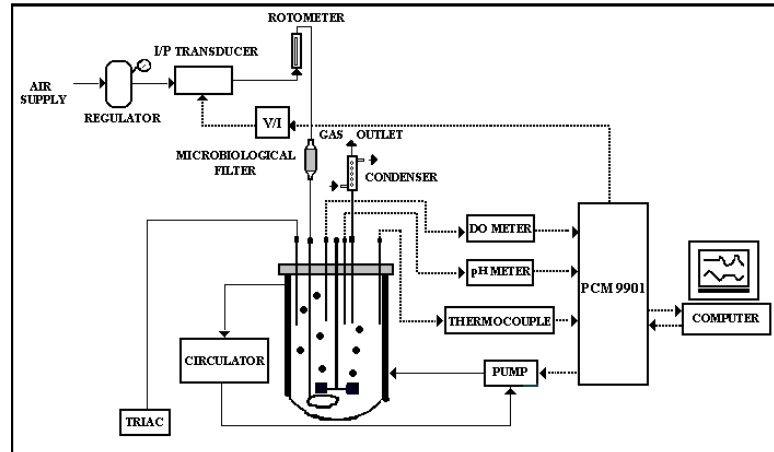


Figure 1. Experimental system

3. Results and Discussion

The discrete-time modelling is very user friendly and the estimation algorithms can be easily implemented utilizing MATLAB. In addition the discrete-time noise processes can be handled in the on-line bioprocess. For bioprocesses, if physics is not well understood although some knowledge about processes is always available, a black-box approach can be utilized. This will lead to some iterative procedure. Then further refinements can be achieved using system identification methods.

The system model in discrete-time domain, ARMAX simulates the system transient behavior for control purposes. The model Eq. 1 relates cooling water flow rate and bioreactor temperature. For using experimental and theoretical square wave input and ternary wave input signals for identification, the bioprocess open loop response is obtained. The periodic square input change is applied as 25-35 magnitude of pump signal. The ternary input change with three values of 25-35-50 also is given as periodic pseudo disturbance. In the face of different periodical effects, temperature transients of bioprocess medium were examined theoretically and experimentally (Figures 2, 3, 4 and 5).

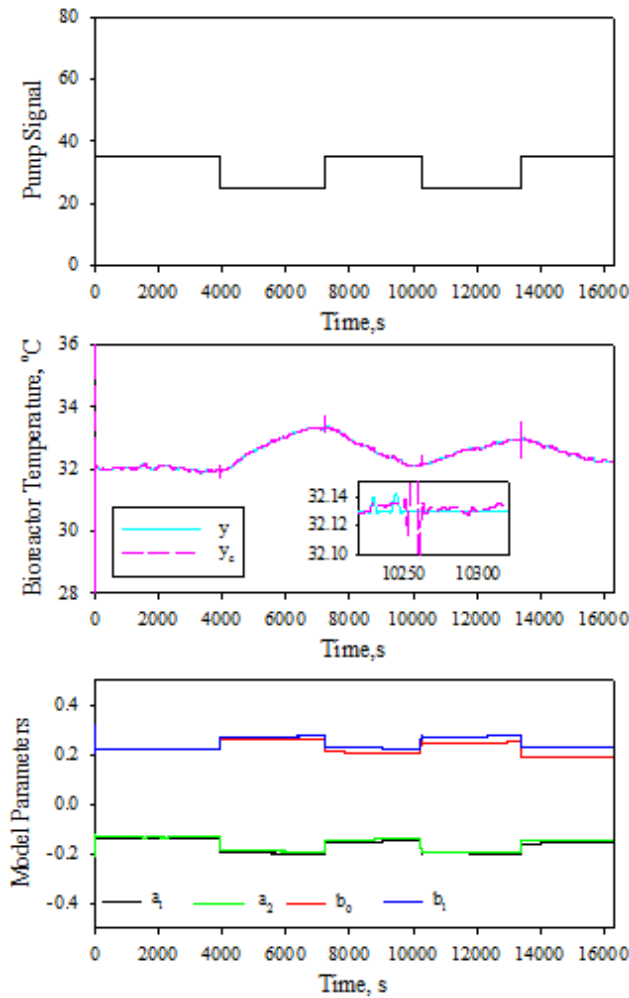


Figure 2. Identification of second order model with square wave input (experimental)

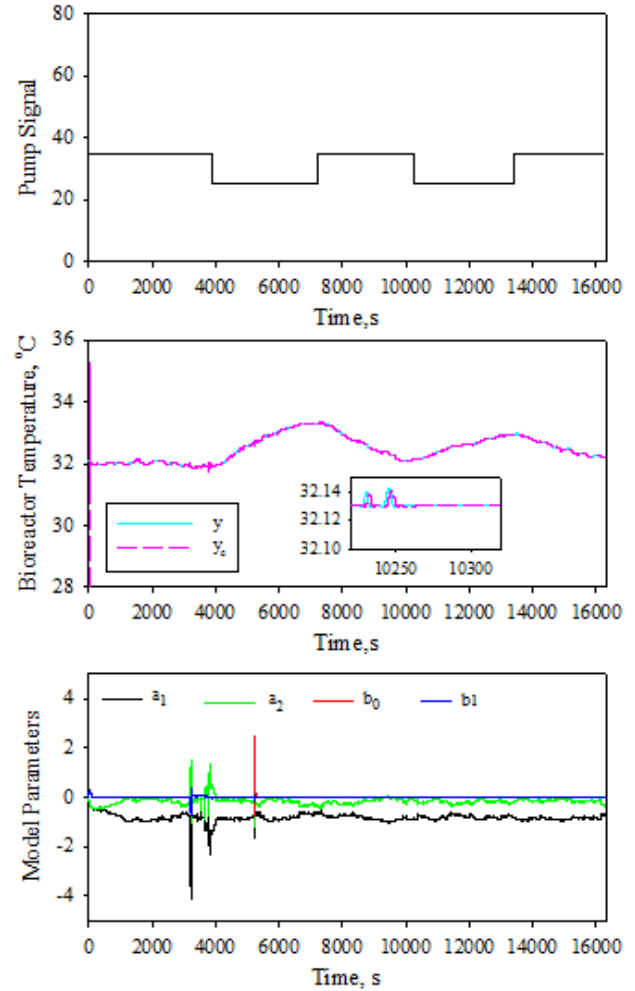


Figure 3. Identification of second order model with square wave input (theoretical)

A RLS technique was utilized in MATLAB package for estimation of various order ARMAX model parameters. These models and their estimated parameters are shown in Table 1. Performance of the second order system models were illustrated in the face of square and ternary external force dynamic behavior in Figure 2, 3 and 4.

Different model order can be viable but over-parameterization cause problems. The third order model for this study can identify the transient behavior of bioprocess medium which is shown in Figure 5.

By means of the recursive least square method, ARMAX model parameter changes were found in the face of the certain pump signal disturbance. The calculated values of ARMAX model parameters are shown in Table 1.

Table 1. Model identification results in the face of the ternary disturbance

| ARMAX model | Evaluation with experimental data | Evaluation with theoretical data |
|--|--|---|
| $y(t)=b_0u(t-1)+b_1u(t-2)$ $-a_1y(t-1)-a_2y(t-2)$ | $a_1=-0.142$ $b_0=0.183$ $a_2=-0.164$ $b_1=0.228$ | $a_1=-0.886$ $b_0=3.26 \times 10^{-7}$ $a_2=-0.114$ $b_1=-8.27 \times 10^{-5}$ |
| $y(t)=b_0u(t-1)+b_1u(t-2)+b_2u(t-3)$ $-a_1y(t-1)-a_2y(t-2)-a_3y(t-3)$ | $a_1=-0.187$ $b_0=0.186$ $a_2=-0.135$ $b_1=0.214$ $a_3=-0.185$ $b_2=0.242$ | $a_1=-0.893$ $b_0=-8.97 \times 10^{-7}$ $a_2=-0.171$ $b_1=-2.47 \times 10^{-4}$ $a_3=0.064$ $b_2=1.69 \times 10^{-4}$ |

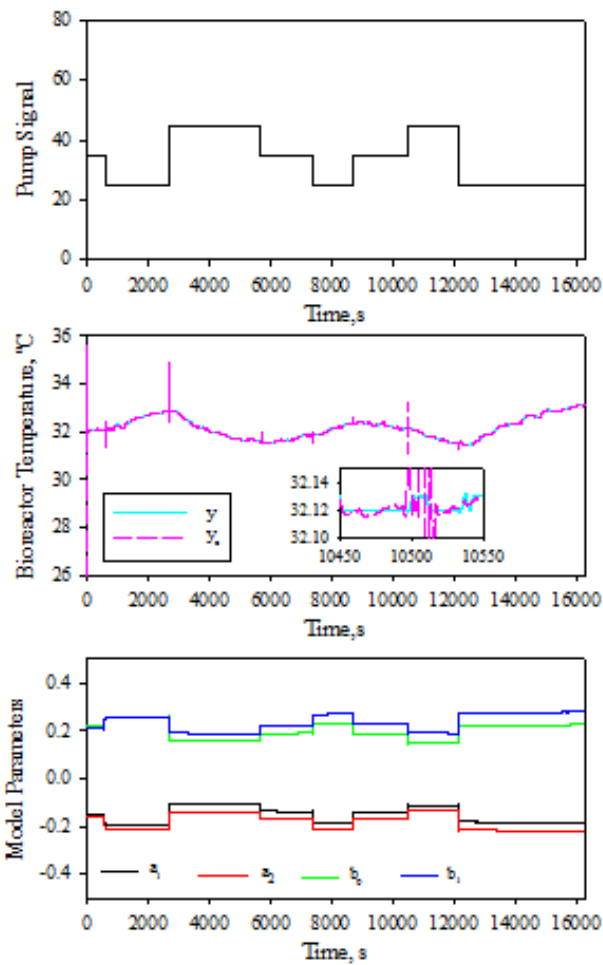


Figure 4. Identification of second order model with ternary wave input (experimental)

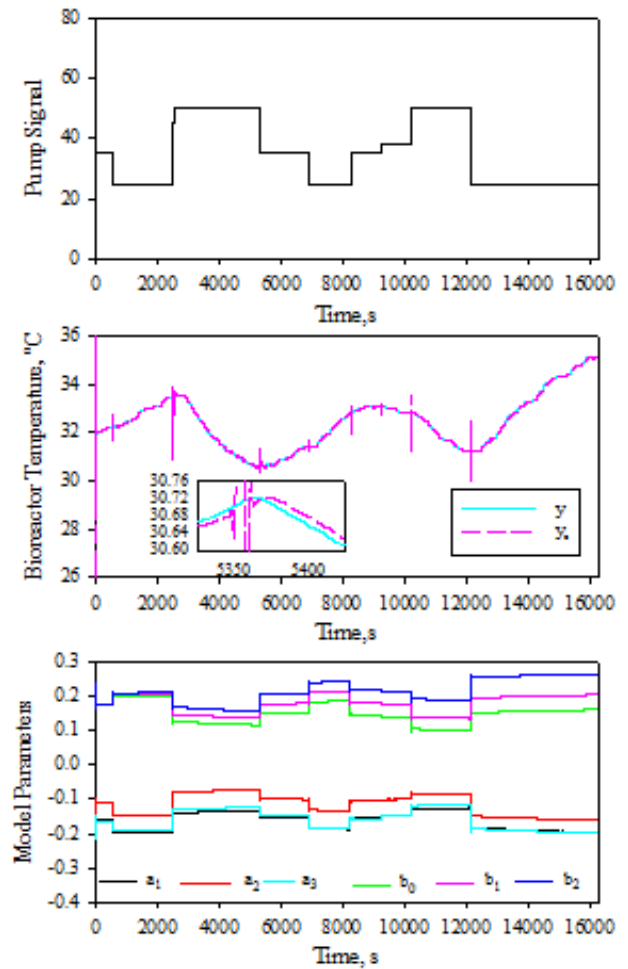


Figure 5. Identification of third order model with ternary wave input (experimental)

4. Conclusion

The identified models given in Table 2 can be utilized in discrete time advanced control algorithm. The various order ARMAX models are tested with evaluation of IAE, ISE and PEEN criteria.

As the model order increases to the true value, the IAE, ISE and PEEN values will decrease. Then a better fit can be obtained by using theoretical data obtained from bioprocess simulation program. In this theoretical study, use of ternary input signal type disturbance

gives better results than square input signal application to obtain system response in second order model identification according to three performance criteria mentioned, (see Table 2). It is noted that considerable experimental errors occur in bioreactor temperature response compare to simulation result as it is expected in experimental cases studied (see Figure 4 and Figure 5). These errors have high impact on three performance criteria calculated in Table 2. Thus, in experimental cases, second and third order model performance should be compared by using estimated and real time bioreactor temperature changes with time in Figure 4 and Figure 5.

Table 2. Performance criteria for ARMAX model with two different orders.

| Input Signal Type | Model Order | EXPERIMENTAL | | | THEORETICAL | | |
|-------------------|-------------|--------------|--------|--------|-------------|--------|--------|
| | | IAE | ISE | PEEN | IAE | ISE | PEEN |
| Square | Second | 23.672 | 1.3876 | 0.0091 | 20.995 | 0.3173 | 0.0081 |
| | Second | 30.102 | 8.977 | 0.0112 | 18.865 | 0.133 | 0.0073 |
| Ternary | Third | 94.410 | 18.497 | 0.0361 | 12.140 | 0.125 | 0.0047 |

Acknowledgements

Authors are thankful to Ankara University Research Fund and Biotechnology Institute (Project Number:29) for the encouragement and support to carry out the research. The Corresponding author also gratefully acknowledged Scientific and Technological Research Council of Turkey (TUBITAK) for PhD National Scholarship Programme.

5. References

- [1]. Akay, B., Ertunç, S., Boyacıoğlu H., Hapoğlu, H., Albaz, M. (2011). Discrete system identification and self-tuning control of dissolved oxygen concentration in a stirred reactor. *Korean J of Chem. Eng.*, 28(3), 837-847.
- [2]. Huang, B., Guo J., Yi, B., Yu, X., Sun, L., Chen, W. (2008). Heterologous production of secondary metabolites as pharmaceuticals in *Saccharomyces cerevisiae*. *Biotechnol Lett.*, 30, 1121-1137.
- [3]. Ertunç, S., Akay B., Bursali N., Hapoğlu H., Albaz M. (2003). Generalized minimum variance control of growth medium temperature of Baker's yeast production, *Food and Bioproducts Processing*, 81, 327-335.
- [4]. Warwick, K., Rees, D. (1988). *Industrial Digital Control Systems*. Peter Peregrinus, London.
- [5]. Svoronos, S., Stephanopoulos, G., Aris, R. (1981). On bilinear estimation and control, *Int J Control*, 34(4), 651-684.
- [6]. Hapoğlu, H., Karacan, S., Koca Erten, Z.S., Albaz, M. (2001). Parametric and nonparametric model based control of a packed distillation column. *Chem. Eng. and Processing*, 40, 537-544.
- [7]. Akay, B., Ertunç, S., Boyacıoğlu, H., Hapoğlu, H., Albaz, M. (2003). *Parametric and non-parametric models' identification based-on dissolved oxygen concentration in S.cerevisiae production*. 3rd Chemical Engineering Conference for Collaborative Research in Eastern Mediterranean (EMCC-3), Thessaloniki (Greece)
- [8]. Zhu, X., Seborg, D.E. (1994). *Non-linear predictive control based on Hammerstein models*. PSE'94 Conference, June, Kyongju, Korea.

- [9]. Norquay, S., Palazoglu, A., Romagnoli, J.A. (1996). *Nonlinear model predictive control of pH neutralization using Wiener models*. Proceedings of the IFAC World Conference, San Francisco, M, 31.
- [10]. Wang, D., Zhou, D.H., Jin, Y.H., Qin, S.J. (2004). Adaptive generic model control for a class of non-linear time varying processes with input time delay. *J. Process Contr.*, 14, 517-531.
- [11]. Sotomayor, O.A.Z., Park, S.W. (2003). Garcia C. Multivariable identification of an activated sludge process with subspace based algorithms. *Contr. Eng. Pract.*, 11: 961-969.
- [12]. Sima, V., Sima, D.M., Van Huffel, S. (2004). High performance numerical algorithms and software for subspace-based linear multivariable system identification. *J. Com. Appl. Math.*, 170, 371-397.
- [13]. Soderström, T., Stoica, P. (1998). *System Identification*. Prentice Hall Ltd. New York.
- [14]. Boyacıoğlu, H. (2013). Çok Girdili-Çok Çıktılı Kontrolün Ekmek Mayası Üretimine Uygulanması. *Doktora Tezi*, Ankara Üniversitesi Fen Bilimleri Enstitüsü, Ankara.